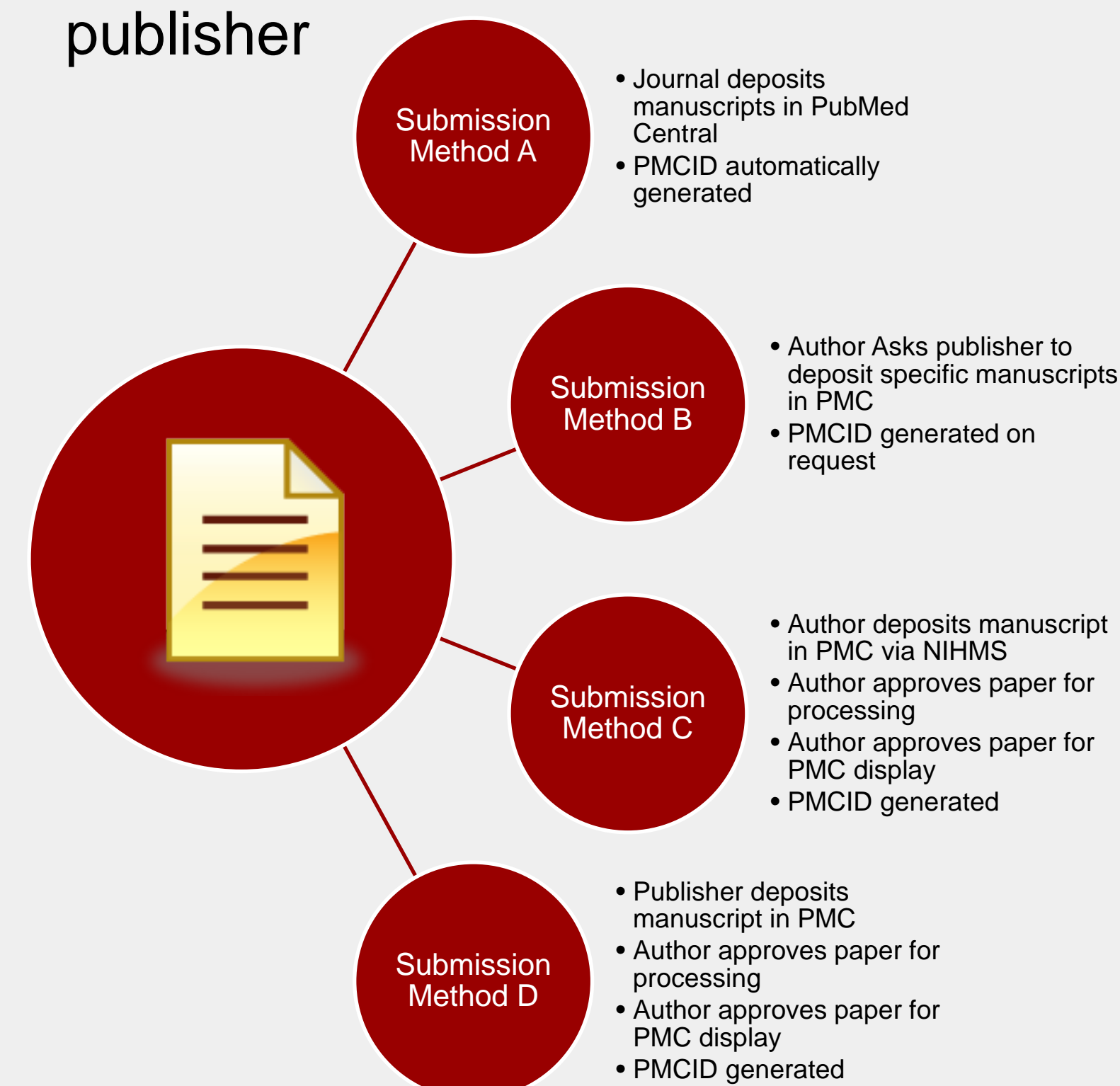


Supporting NIH Public Access Policy compliance using VIVO data

Drew Wright* and Paul Albert
Weill Cornell Medical Library, New York, NY

The Problem

- The NIH has begun strictly enforcing their Public Access Policy, which requires all funded publications be made publically available within twelve months of publication
- Few authors keep track of or are even aware of the policy
- Submission methods and necessary steps are confusing and vary by publisher



The Solution?

- Coupling VIVO data with authoritative NIH data, it is possible to identify non-compliant publications and notify grant holders and authors of necessary steps

Goals and Objectives

- Construct SPARQL queries to identify non-compliant papers in VIVO instance
- Develop quality control and notification workflows with little manual oversight

Contact

Drew Wright
drw2004@med.cornell.edu

Downloads

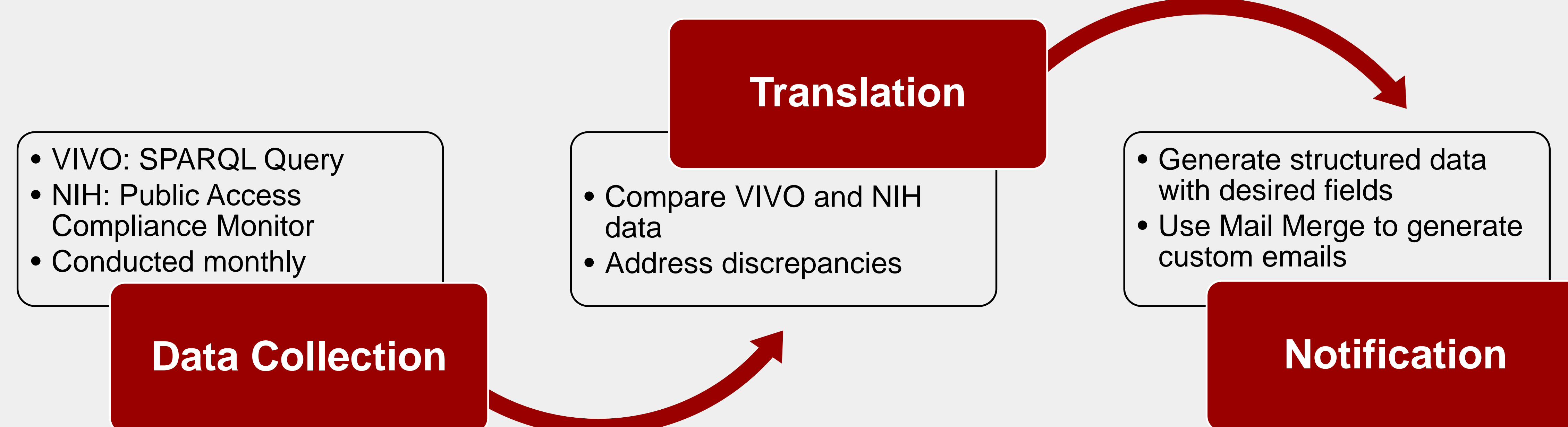


Poster



Sample Data

Process Workflow



- VIVO: SPARQL Query
- NIH: Public Access Compliance Monitor
- Conducted monthly

Data Collection

SPARQL Query

```
SELECT distinct ?Person1_firstName
?Person1_lastName ?Person1_primaryEmail
?AcademicArticle1_label ?Journal1_label
?AcademicArticle1_pmid
?DateTimeValue1_dateTime

WHERE{
?AcademicArticle1 rdf:type bibo:Document .
?AcademicArticle1 bibo:pmid .
?AcademicArticle1_pmid .
?AcademicArticle1_vivo:dateTimeValue
?DateTimeValue1 .
?AcademicArticle1
vivo:informationResourceSupportedBy
?FundingOrganization1 .
?AcademicArticle1 bibo:pmid .
?AcademicArticle1_pmid .
?DateTimeValue1 rdf:type vivo:DateTimeValue .
?DateTimeValue1_vivo:dateTime
?DateTimeValue1_dateTime .
?FundingOrganization1 rdf:type
vivo:FundingOrganization
?FundingOrganization1_rdfs:label
?FundingOrganization1_label .
?AcademicArticle1_rdfs:label
?AcademicArticle1_label .
?AcademicArticle1_vivo:hasPublicationVenue
?Journal1 .
?Journal1 rdf:type bibo:Journal .
?Journal1_rdfs:label ?Journal1_label .
?AcademicArticle1
vivo:informationResourceInAuthorship ?Authorship1
?Authorship1 rdf:type vivo:Authorship .
?Authorship1_vivo:linkedAuthor ?Person1 .
?Person1 rdf:type foaf:Person .
?Person1_vivo:primaryEmail
?Person1_primaryEmail .
?Person1_wcmc:cwid ?Person1_cwid .
?Person1 foaf:firstName ?Person1_firstName
?Person1 foaf:lastName ?Person1_lastName
FILTER REGEX (str(?FundingOrganization1_label),
"NIH");
FILTER NOT EXISTS { ?AcademicArticle1
vivo:pmcid ?AcademicArticle1_pmid . }
FILTER (xsd:dateTime(?DateTimeValue1_dateTime)
> "2008-04-01T00:00:00"^^xsd:dateTime)
FILTER (xsd:dateTime(?DateTimeValue1_dateTime)
< "2012-12-01T00:00:00"^^xsd:dateTime)
}
```

Returns all publications
in VIVO that are indexed
as having NIH funding
but do not have a PMCID

Translation

- Compare VIVO and NIH data
- Address discrepancies

Notification

- Generate structured data with desired fields
- Use Mail Merge to generate custom emails

Notification Template

Dear Dr. <<Last Name>>,

You are receiving this message because you are listed as PI affiliated with a grant listed on an article that is not in compliance with NIH's Public Access Policy:

"<<Article Title>>" published in <<Journal>> and linked to grant number <<Grant ID>>. Link: <http://www.ncbi.nlm.nih.gov/pubmed/<<PMID>>>

NIH's Public Access Policy requires that all manuscripts resulting from NIH funding be deposited in PubMed Central within three months of publication. Failure to comply could negatively affect the renewal of any associated grants or future grant proposals. It is highly recommended that you take the appropriate steps to ensure compliance as soon as possible. We are available for individual or group consultations to assist you with this process, and please feel free to schedule an appointment with a member of our research support team at your convenience by emailing infodesk@med.cornell.edu or responding to this email. Additionally, below are a few helpful resources that should answer any pressing questions.

Weill Cornell Medical Library NIH Public Access Policy Guide - includes links to online video tutorials and compliance instructions

NIH Manuscript Submission System Help - includes robust FAQ, video tutorials, and step-by-step deposition instructions

Please let me know if you have any additional questions or if you have received this notification in error (not your grant, grant was not affiliated with paper, paper does not fall under the policy, etc - for those of you who have not had the pleasure of interacting with the NIH system, the quality of the data is severely lacking, but we're working on cleaning it up). Please ignore if you have already taken steps toward compliance.

From VIVO

- Author name
- Author email
- PMID

From NIH

- PI name
- Grant number
- PMID

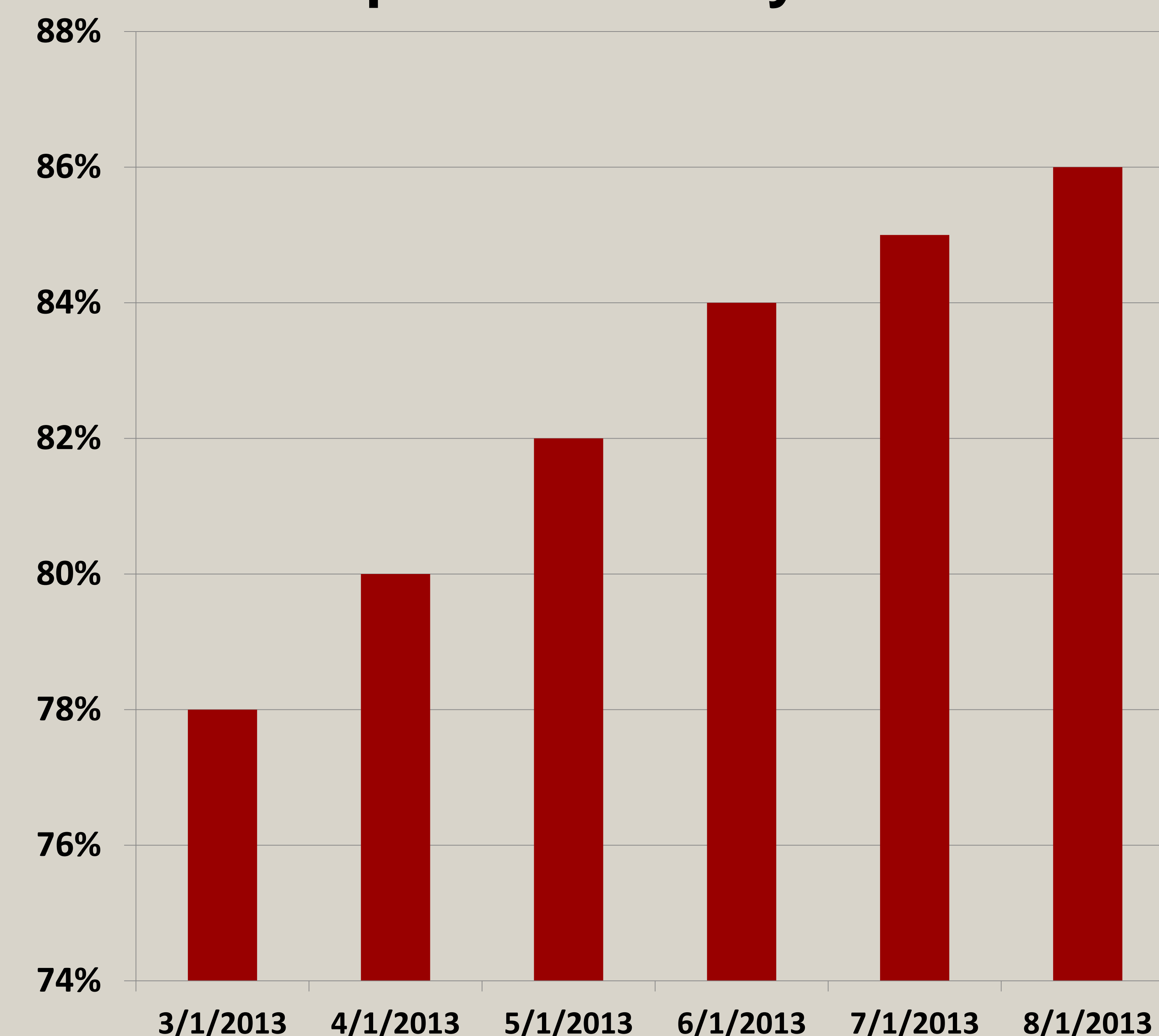
Publications in VIVO but not in NIH (false positives)

- Papers resulting in funding not directly affiliated with WCMC, but with WCMC authors
- Papers incorrectly indexed as resulting from WCMC grants
- Papers accepted before policy date, but published afterwards

Publications in NIH but not in VIVO (false negatives)

- Papers were not ingested completely or at all
- Authors no longer affiliated with WCMC (graduate students, faculty who have moved on, etc.)

Compliance rate by month



Discussion

- First round of notifications sent in March 2013, compliance rate has increased from 78% to 86% (compliance was consistently below 80% prior to intervention)
- Large amount of false positives requires manual quality control, though this is getting more manageable with each iteration
- We target corresponding authors when possible, but affiliation data is not always available
- Publications that contain multiple grants from multiple PIs or are associated with training grants, multi-center awards, or the CTSA program represent a unique challenge since who exactly is responsible for ensuring compliance is not concrete

Future Work

- Fully automate identification and notification:
 - Schedule SPARQL queries
 - Automate exporting of NIH data
 - Identify and permanently exclude false positives
- Ingest corresponding author affiliation and contact information from Scopus, which will allow us to contact non-WCMC affiliated authors directly
- Develop a way to integrate separate workflows for different groups or funding types
 - This has been done to a certain extent for the WCMC CTSC, but requires additional manual work